REMARKS

Claims 5-8, 11, and 12-19 are rejected under 35 U.S.C. §101 because the claimed invention is directed to non-statutory subject matter. Claims 5, 6 and 12 are amended to overcome this rejection. Claim 5is amended to describe the search method to be part of or in a recognizer processing device in a speech recognition device having the recognizer processing device and an output device for presenting the recognized speech to a user. Claim 5, as amended, is therefore deemed to cover patentable subject matter. Claim 6, as amended, describes the method of decoding multiple HMM sets to be in a speech recognition device including a recognizer processing device and an output device for presenting the recognized speech to a user. Claim 6, as amended, is therefore deemed to cover patentable subject matter. Claim 12 describes the means for decoding a plurality of model sets using a generic base grammar network to be in a recognizer processing device that is part of a speech recognizer including the recognizer processing device and an output device for presenting the recognized speech to a user. Claim 12, as amended, is therefore deemed to cover patentable subject matter. Clearly the amended claims call for being part of a speech recognizer with post solution activity and thereby making these claims statutory subject matter. The claims 7-8 and 13-19 are dependent on the amended claims and are deemed statutory subject matter for at least the same reasons.

Claims 3, 5, 12-16, and 19 are rejected under 35 U.S.C. § 102 (e) as being anticipated by Neumeyer et al. (U.S. patent No. 6,226,611; hereinafter Neumeyer).

Applicant believes that this was adequately addressed in the last response but the examiner does not appear to understand. Applicant's previous remarks in the paper sent on March 31, 2006 are incorporated herein. The examiner does not tell applicant why the

arguments put forth in our previous paper are not sufficient proof of the novelty of applicant's claimed invention. The references that the examiner cites nowhere cover or teach the novel steps implemented in applicant's recognizer. The examiner appears to need the definition of "set of HMMs" in the claims to distinguish from a plurality of HMMs so applicant has amended the claims to state what is in the specification that "each HMM set of said HMM sets is a group of HMMs from one environment."

In regard to the detailed action item (4), it appears that the main issue is the examiner's misunderstanding of the term "sets of HMMs" and a "generic grammar network." The examiner states on page 6 of the Office Action, "Neumeyer, et al., teach or discuss a plurality, or sets of HMMs, i.e., a network."

A plurality of HMMs is not the same as a "set of HMMs" in applicant's patent application. In applicant's patent application a "set of HMMs" is a group of HMMs **from one environment**. Here, an environment may be male speech, or female speech, or may include any other such divisions of speech such as dialects or accents. This is an important distinction, because applicant's method of speech recognition will recognize an utterance using multiple sets of HMMs, wherein the recognition method is constrained to calculate the utterance likelihood from each set of HMMs without using HMMs from any other set. That is, applicant's recognizer method determines the most likely utterance, which may be from the male environment, or the female environment, etc., but there is no mixing of environments. It does this in a novel manner that does not require that the grammar exhaustively enumerate all valid sequences of each model set. The Neumeyer reference clearly does not address this distinction of environments.

A plurality of HMMs, or set of HMMs, is not a network. HMMs are models of a word or sub-word units, such as phone models. A network is a structure that defines the valid sequences of HMMs that the recognizer is allowed to process to determine the most likely valid sequence. For example, suppose that a set of HMMs consisted of the two words "yes" and "no". Then one grammar network may allow either the word "yes" or the word "no" as valid recognition sequence. On the other hand, another grammar network may allow only "yes yes no" or "no yes no" as valid recognition sequences. Clearly, a plurality, or set of HMMs, is not synonymous with a network.

This is also an important distinction, since in applicant's recognition method it is required that the recognizer constrain its valid sets of sequences to within an HMM set. For example, suppose we had two sets of HMMs for the word " yes" and "no", one for the male environment, and one for the female environment. Thus we would have two sets of HMMs. The set for the male would consist of the HMMs "yes:m" and "no:m", where the ":m" indicates these models correspond to the male environment. The set for the female would consist of the HMMs "yes:f" and "no:f" , where the ":f" indicates these models correspond to the female environment. Now suppose that we want a recognizer to recognize only the two word sequences "yes yes no" or "no yes no". In the multiple environment case, we want to only allow "yes:m yes:m no:m", or "no:m yes:n no:m", or "yes:f yes:f no:f", or "no:f yes:f no:f". Note that the recognizer is not allowed to consider any sequence where the male and female HMMs are intermixed, such as "yes:m yes:f no:m". The present art method of doing this is to create a single grammar network that enumerates all possible sequences of HMMs. Thus the present art grammar network would consist of:

"yes:m yes:m no:m"

"no:m yes:m no:m"

"yes:f yes:f no:f"

"no:f yes:f no:f"

The method of applicant's invention is to use a "generic grammar network", which is defined as a network that does not specify the set-dependent HMMs as the valid sequences. Thus the generic grammar network consists of only:

"yes yes no"

"no yes no".

In accordance with applicant's present invention, applicant's teach processing steps within the speech recognition processor to map the words of the generic grammar network (yes, no) to the individual sets of HMMs and to keep track of speech recognition likelihood processing so that likelihoods for separate HMM sets are not intermixed, in order to have the resulting recognizer processor output the same results as a speech recognizer using the present art grammar network above in which all allowable set-dependent HMM sequences must be enumerated. This results in memory savings and processing step efficiency.

It is clear that the Neumeyer reference does not discuss or teach the concept of HMM set-dependent networks and using a generic grammar network and mapping steps to implement an HMM set-dependent speech recognition processor by only using a generic grammar network.

In view of the above applicant's claims 3, 5, 12-16, and 19 are not taught in Neumeyer and are improperly rejected under 35 U.S.C. 102 (e).

Claim 6 is rejected under 35 U.S.C. § 102(b) as being anticipated by Naylor et al. (U.S. Patent No. 5,806,033; hereinafter Naylor).

The examiner states, "Naylor teaches multiple sets of HMMs because of the separate use of male speakers to train the HMMs (Col. 6, lines 15-25) leading one to naturally conclude that a separate set of trained HMMs (using the male speakers) is available along with a set of trained HMMs using a general population." In this statement, the examiner correctly points out that Naylor does mention collecting data from the male environment. It might even be a valid argument that one skilled in the art might deduce that it would be possible to create a set of HMM models or even a set of female HMM models. However, applicant does not claim this as the novel inventive steps of the application. Actually, such creation of sets of HMMs environment models is well known in the present art. What applicant claim is a method of speech recognition for decoding multiple HMM sets according to the inventive steps of our patent application.

With respect to claim 6 the inventive steps are providing a generic network containing base symbols; a plurality of sets of HMMs where each set of HMMs corresponds to a single environmental factor such as for male and female; each said set of HMMs enumerated in terms of expanded symbols which map to the generic network base symbols; accessing said generic network using said base symbols through a conversion function that gives base symbols for expanded symbols to therefore decode multiple HMM sets using a generic base sentence grammar and using said HMM sets to recognize incoming speech.

Naylor does not teach such processing.

Considering each of the arguments by the examiner they do not support the examiner's position.

The examiner states, "...Naylor et al., teach a method of speech recognition for decoding multiple HMM sets using a generic base sentence network comprising the steps of: providing a generic network containing base symbols (Col. 3, lines 30-40, Fig. 2, items 32,34,36);" However, Col. 3, lines 30-40 fail to make any mention of any method of decoding or any grammar network, and Fig, 2 is a figure that describes how HMMs are trained prior to recognition, and hence there can be no decoding of HMM sets in items

32,34,and 36. Items 32, 34, and 36 just depict preparation of speech databases and initialization of HMM structures prior to training HMMs. Note also that generic network base symbols do not correspond one-to-one with HMMs, but rather to each generic network base symbol there correspond HMMs from each of the HMM sets.

The examiner further states, "...each said set of HMMs enumerated in terms of expanded symbols which map to the generic network base symbols (Fig. 4);" However, Fig. 4, is an overview flow diagram of the well-known Viterbi decoding algorithm for HMM-based speech recognition. Nowhere in Fig. 4 is there any processing step that utilizes a generic network of base symbols and processing to enumerate expanded symbols that map to base symbols.

The examiner further states,"...accessing said generic network using said base symbols through a conversion function that gives base symbols for expanded symbols to therefore decode multiple HMM sets using a generic base sentence grammar and using said HMM sets to recognize incoming speech (Figs. 5-7, Col. 8, lines 45-52, Col. 7, lines 49-55). However, Fig. 5 is a flow diagram of the well-known Viterbi forward likelihood computation algorithm for propagating state likelihoods for a given input frame, and managing back pointers to the prior best-likelihood state. This operation does not involve any grammar network processing, since the processing occurs within an HMM itself, and back propagation is to an HMM state.

Fig. 6 is a flow diagram of concatenating symbol strings to determine decoded words. However, this does not mention anywhere the use of a grammar network of base symbols to enumerate expanded symbols that map to base symbols during recognition in order to recognize the multiple HMM sets enumerated by the expanded symbols, maintaining the separate likelihoods for each HMM set. Indeed, in Figures 5 and 6 it is clear that no such separation of likelihoods of expanded symbols is taught or suggested.

Fig. 7 is a diagram of a backtracking processing step well-known in the art. This diagram does not show or teach using a generic grammar network of base symbols to enumerate expanded symbols that map to the base symbols. If it did, this would be clearly demarked in the region of 94 of the diagram as pertaining to multiple HMM sets.

Col. 8, lines 45-52 simply describe overall processing shown in Figures 5-7. Again, it nowhere mentions generic network using said base symbols through a conversion

function that gives base symbols for expanded symbols to therefore decode multiple HMM sets using a generic base sentence grammar and using said HMM sets to recognize incoming speech.

Col. 7, lines 49-55 state that a grammar is used to enumerate valid inter-word transitions. However, this does not mention that the grammar is a generic network that incorporates a mapping conversion function to decode multiple HMM sets which are specified by the expanded symbols that map to the base symbols.

It is clear to one skilled in the art that the Naylor reference teaches utilizing a finite state grammar method (Col. 9, lines 61-64) well-known in the art to enumerate all possible symbol sequences, without using a grammar consisting of base symbols to enumerate expanded symbols that map to the base symbols. It is further clear that Naylor does not teach or describe a method to utilize any such grammar consisting of base symbols and a mapping function to enumerate expanded symbols corresponding to HMM sets to separately determine the probability of sequences of HMMs in which each valid sequence is composed of HMMs from one set only.

In view of the above applicants Claims 3, 5-8 and 10-19, as amended, are deemed allowable and an early notice of allowance of these claims is deemed in order and is respectfully requested.

Respectfully requested;

Robert L. Troike (Reg. 24183)

Telephone No. (301) 751-0825